Algorithm 1 BiGAN Model

1: **Input Batch**: images x, correct text caption t, wrong text caption $\hat{t}$, total batch training steps N
2: **for** i=1 to N **do**
3:      $e \leftarrow Encoder(x)$ $\triangleright$ Convolving the image
4:      $h \leftarrow \phi(t)$ $\triangleright$ Encode correct text caption
5:      $\hat{h} \leftarrow \phi(\hat{t})$ $\triangleright$ Encode wrong text caption
6:      $z \leftarrow \mathcal{N}(0,1)$ $\triangleright$ Generate gaussian noise
7:      $\hat{x} \leftarrow G(z,h,e)$ $\triangleright$ Generator forward pass
8:      $l \leftarrow \phi(c)$ $\triangleright$ Encode correct logo code
9:      $\hat{c} \leftarrow R(D(x))$ $\triangleright$ Generate logo code
10:      $\hat{l} \leftarrow \phi(\hat{c})$ $\triangleright$ Encode generated logo code
11:      $d_1 \leftarrow D(x,h)$ $\triangleright$ Real image and correct caption
12:      $d_2 \leftarrow D(x,\hat{h})$ $\triangleright$ Real image and wrong caption
13:      $d_3 \leftarrow D(\hat{x},h)$ $\triangleright$ Fake image and correct caption
14:      $L_D \leftarrow \log(d_1) + \log(1-d_2)/2 + \log(1-d_3)/2$ $\triangleright$ Discriminator Loss
15:      $D \leftarrow D - \alpha \partial L_D/\partial D$ $\triangleright$ Update Discriminator Parameters
16:      $L_E \leftarrow |x - \hat{x}|$ $\triangleright$ Encoder L1 Loss
17:      $L_G \leftarrow \log(d_3) + L_E$ $\triangleright$ Generator Loss
18:      $L_R \leftarrow \|l, \hat{l}\|$ $\triangleright$ RNN Loss
19:      $G \leftarrow G - \alpha \partial L_G/\partial G$ $\triangleright$ Update Generator Parameters
20:      $R \leftarrow R - \alpha \partial L_R/\partial R$ $\triangleright$ Update RNN Parameters
21: **end for**